(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 318 216 B1

# (12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: 15.12.93

(51) Int. Cl.5: C12N 15/00, C12N 7/00, A61K 39/29, G01N 33/576, C12Q 1/70, C12Q 1/68, C07K 15/00, A61K 39/395, C12P 21/00

(21) Application number: 88310922.5

(22) Date of filing: 18.11.88

The file contains technical information submitted after the application was filed and not included in this specification

## (54) NANBV diagnostics and vaccines.

(30) Priority: 18.11.87 US 122714
30.12.87 US 139886
26.02.88 US 161072
06.05.88 US 191263
26.10.88 US 263584
14.11.88 US 271450

(43) Date of publication of application: 31.05.89 Bulletin 89/22

(45) Publication of the grant of the patent: 15.12.93 Bulletin 93/50

(84) Designated Contracting States: AT BE CH DE ES FR GR IT LI LU NL SE

(56) References cited:
EP-A- 0 194 207
EP-A- 0 279 460
WO-A-82/02774
EP-A- 0 263 761
WO-A-82/00205
WO-A-88/03410

(73) Proprietor: CHIRON CORPORATION
4560 Horton Street
Emeryville, California 94608(US)

(72) Inventor: Houghton, Michael
1307 Greenbrook Drive
Danville California 94526(US)
Inventor: Choo, Qui-Lim
5700 Fern Street
El Cerrito California 94530(US)
Inventor: Kuo, George
1370 Sixth Avenue
San Francisco California 94122(US)

(74) Representative: Goldin, Douglas Michael et al
J.A. KEMP & CO.
14, South Square
Gray's Inn
London WC1R 5EU (GB)



EP 0 318 216 B1

Rank Xerox (UK) Business Services
(3.10/3.6/3.3.1)

sequence of the polypeptide encoded in the extended ORF in the derived sequence.

Fig. 27 shows the sequence of the HCV cDNA in clone 12f, the segment which overlaps clone 14i, and the amino acids encoded therein.

Fig. 28 shows the sequence of the HCV cDNA in clone 35f, the segment which overlaps clone 39c, and the amino acids encoded therein.

Fig. 29 shows the sequence of the HCV cDNA in clone 19g, the segment which overlaps clone 35f, and the amino acids encoded therein.

Fig. 30 shows the sequence of clone 26g, the segment which overlaps clone 19g, and the amino acids encoded therein.

Fig. 31 shows the sequence of clone 15e, the segment which overlaps clone 26g, and the amino acids encoded therein.

Fig. 32 shows the sequence in a composite cDNA, which was derived by aligning clones 12f through 15e in the 5' to 3' direction; it also shows the amino acids encoded in the continuous ORF.

Fig. 33 shows a photograph of Western blots of a fusion protein, SOD-NANB$_{5-1-1}$, with chimpanzee serum from chimpanzees infected with BB-NANB, HAV, and HBV.

Fig. 34 shows a photograph of Western blots of a fusion protein, SOD-NANB$_{5-1-1}$, with serum from humans infected with NANBV, HAV, HBV, and from control humans.

Fig. 35 is a map showing the significant features of the vector pAB24.

Fig. 36 shows the putative amino acid sequence of the carboxy-terminus of the fusion polypeptide C100-3 and the nucleotide sequence encoding it.

Fig. 37A is a photograph of a coomassie blue stained polyacrylamide gel which identifies C100-3 expressed in yeast.

Fig. 37B shows a Western blot of C100-3 with serum from a NANBV infected human.

Fig. 38 shows an autoradiograph of a Northern blot of RNA isolated from the liver of a BB-NANBV infected chimpanzee, probed with BB-NANBV cDNA of clone 81.

Fig. 39 shows an autoradiograph of NANBV nucleic acid treated with RNase A or DNase I, and probed with BB-NANBV cDNA of clone 81.

Fig. 40 shows an autoradiograph of nucleic acids extracted from NANBV particles captured from infected plasma with anti-NANB$_{5-1-1}$, and probed with $^{32}$P-labeled NANBV cDNA from clone 81.

Fig. 41a and b shows autoradiographs of filters containing isolated NANBV nucleic acids, probed with $^{32}$P-labeled plus and minus strand DNA probes derived from NANBV cDNA in clone 81.

Fig. 41-1 shows the homologies between a polypeptide encoded in HCV cDNA and an NS protein from Dengue flavivirus.

Fig. 43 shows a histogram of the distribution of HCV infection in random samples, as determined by an ELISA screening.

Fig. 44 shows a histogram of the distribution of HCV infection in random samples using two configurations of immunoglobulin-enzyme conjugate in an ELISA assay.

Fig. 45 shows the sequences in a primer mix, derived from a conserved sequence in NS1 of flaviviruses.

Fig. 46 shows the HCV cDNA sequence in clone k9-1, the segment which overlaps the cDNA in Fig. 27, and the amino acids encoded therein.

Fig. 47 shows the sequence in a composite CDNA which was derived by aligning clones k9-1 through 15e in the 5' to 3' direction; it also shows the amino acids encoded in the continuous ORF.

## I. Definitions

The term "hepatitis C virus" has been reserved by workers in the field for an heretofore unknown etiologic agent of NANBH. Accordingly, as used herein, "hepatitis C virus" (HCV) refers to an agent causitive of NANBH, which agent is a virus characterised by:(i) a positive stranded RNA genome; (ii) said genome comprising an open reading frame (ORF) encoding a polyprotein; and (iii) the portion of said polyprotein corresponding to Figure 14 having at least 40% homology to the amino acid sequence in Figure 14. This agent was formerly referred to as NANBV and/or BB-NANBV. The terms HCV, NANBV, and BB-NANBV are used interchangeably herein, but all refer to the virus as defined above. As an extension of this terminology, the disease caused by HCV, formerly Called NANB hepatitis (NANBH), is called hepatitis C. The terms NANBH and hepatitis C may be used interchangeably herein.

The term "HCV", as used herein, denotes a viral species which causes NANBH, and attenuated strains or defective interfering particles derived therefrom. As shown infra., the HCV genome is comprised of RNA. It is known that RNA containing viruses have relatively high rates of spontaneous mutation, i.e., reportedly

FIG. 32-1 COMBINED ORF OF DNAs 12f through 15e

IlePheLysIleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsn
1 CCATATTTAAAATCAGGATGTACGTGGGAGGGGTCGAACACAGGCTGGAAGCTGCCTGCA
GGTATAAATTTTAGTCCTACATGCACCCTCCCCAGCTTGTGTCCGACCTTCGACGGACGT

TrpThrArgGlyGluArgCysAspLeuGluAspArgAspArgSerGluLeuSerProLeu
61 ACTGGACGCGGGGCGAACGTTGCGATCTGGAAGACAGGGACAGGTCCGAGCTCAGCCCGT
TGACCTGCGCCCCGCTTGCAACGCTAGACCTTCTGTCCCTGTCCAGGCTCGAGTCGGGCA

LeuLeuThrThrThrGlnTrpGlnValLeuProCysSerPheThrThrLeuProAlaLeu
121 TACTGCTGACCACTACACAGTGGCAGGTCCTCCCGTGTTCCTTCACAACCCTACCAGCCT
ATGACGACTGGTGATGTGTCACCGTCCAGGAGGGCACAAGGAAGTGTTGGGATGGTCGGA

SerThrGlyLeuIleHisLeuHisGlnAsnIleValAspValGlnTyrLeuTyrGlyVal
181 TGTCCACCGGCCTCATCCACCTCCACCAGAACATTGTGGACGTGCAGTACTTGTACGGGG
ACAGGTGGCCGGAGTAGGTGGAGGTGGTCTTGTAACACCTGCACGTCATGAACATGCCCC

GlySerSerIleAlaSerTrpAlaIleLysTrpGluTyrValValLeuLeuPheLeuLeu
241 TGGGGTCAAGCATCGCGTCCTGGGCCATTAAGTGGGAGTACGTCGTTCTCCTGTTCCTTC
ACCCCAGTTCGTAGCGCAGGACCCGGTAATTCACCCTCATGCAGCAAGAGGACAAGGAAG

LeuAlaAspAlaArgValCysSerCysLeuTrpMetMetLeuLeuIleSerGlnAlaGlu
301 TGCTTGCAGACGCGCGCGTCTGCTCCTGCTTGTGGATGATGCTACTCATATCCCAAGCGG
ACGAACGTCTGCGCGCGCAGACGAGGACGAACACCTACTACGATGAGTATAGGGTTCGCC

AlaAlaLeuGluAsnLeuValIleLeuAsnAlaAlaSerLeuAlaGlyThrHisGlyLeu
361 AGGCGGCTTTGGAGAACCTCGTAATACTTAATGCAGCATCCCTGGCCGGGACGCACGGTC
TCCGCCGAAACCTCTTGGAGCATTATGAATTACGTCGTAGGGACCGGCCCTGCGTGCCAG

ValSerPheLeuValPhePheCysPheAlaTrpTyrLeuLysGlyLysTrpValProGly
421 TTGTATCCTTCCTCGTGTTCTTCTGCTTTGCATGGTATTTGAAGGGTAAGTGGGTGCCCG
AACATAGGAAGGAGCACAAGAAGACGAAACGTACCATAAACTTCCCATTCACCCACGGGC

AlaValTyrThrPheTyrGlyMetTrpProLeuLeuLeuLeuLeuLeuAlaLeuProGln
481 GAGCGGTCTACACCTTCTACGGGATGTGGCCTCTCCTCCTGCTCCTGTTGGCGTTGCCCC
CTCGCCAGATGTGGAAGATGCCCTACACCGGAGAGGAGGACGAGGACAACCGCAACGGGG

ArgAlaTyrAlaLeuAspThrGluValAlaAlaSerCysGlyGlyValValLeuValGly
541 AGCGGGCGTACGCGCTGGACACGGAGGTGGCCGCGTCGTGTGGCGGTGTTGTTCTCGTCG
TCGCCCGCATGCGCGACCTGTGCCTCCACCGGCGCAGCACACCGCCACAACAAGAGCAGC

LeuMetAlaLeuThrLeuSerProTyrTyrLysArgTyrIleSerTrpCysLeuTrpTrp
601 GGTTGATGGCGCTGACTCTGTCACCATATTACAAGCGCTATATCAGCTGGTGCTTGTGGT
CCAACTACCGCGACTGAGACAGTGGTATAATGTTCGCGATATAGTCGACCACGAACACCA

LeuGlnTyrPheLeuThrArgValGluAlaGlnLeuHisValTrpIleProProLeuAsn
661 GGCTTCAGTATTTTCTGACCAGAGTGGAAGCGCAACTGCACGTGTGGATTCCCCCCCTCA
CCGAAGTCATAAAAGACTGGTCTCACCTTCGCGTTGACGTGCACACCTAAGGGGGGGAGT

ValArgGlyGlyArgAspAlaValIleLeuLeuMetCysAlaValHisProThrLeuVal
721 ACGTCCGAGGGGGGCGCGACGCCGTCATCTTACTCATGTGTGCTGTACACCCGACTCTGG
TGCAGGCTCCCCCCGCGCTGCGGCAGTAGAATGAGTACACACGACATGTGGGCTGAGACC

PheAspIleThrLysLeuLeuLeuAlaValPheGlyProLeuTrpIleLeuGlnAlaSer
781 TATTTGACATCACCAAATTGCTGCTGGCCGTCTTCGGACCCCTTTGGATTCTTCAAGCCA
ATAAACTGTAGTGGTTTAACGACGACCGGCAGAAGCCTGGGGAAACCTAAGAAGTTCGGT

LeuLeuLysValProTyrPheValArgValGlnGlyLeuLeuArgPheCysAlaLeuAla
841 GTTTGCTTAAAGTACCCTACTTTGTGCGCGTCCAAGGCCTTCTCCGGTTCTGCGCGTTAG
CAAACGAATTTCATGGGATGAAACACGCGCAGGTTCCGGAAGAGGCCAAGACGCGCAATC

ArgLysMetIleGlyGlyHisTyrValGlnMetValIleIleLysLeuGlyAlaLeuThr
901 CGCGGAAGATGATCGGAGGCCATTACGTGCAAATGGTCATCATTAAGTTAGGGGCGCTTA
GCGCCTTCTACTAGCCTCCGGTAATGCACGTTTACCAGTAGTAATTCAATCCCCGCGAAT

```
     GlyThrTyrValTyrAsnHisLeuThrProLeuArgAspTrpAlaHisAsnGlyLeuArg
961  CTGGCACCTATGTTTATAACCATCTCACTCCTCTTCGGGACTGGGCGCACAACGGCTTGC
     GACCGTGGATACAAATATTGGTAGAGTGAGGAGAAGCCCTGACCCGCGTGTTGCCGAACG

     AspLeuAlaValAlaValGluProValValPheSerGlnMetGluThrLysLeuIleThr
1021 GAGATCTGGCCGTGGCTGTAGAGCCAGTCGTCTTCTCCCAAATGGAGACCAAGCTCATCA
     CTCTAGACCGGCACCGACATCTCGGTCAGCAGAAGAGGGTTTACCTCTGGTTCGAGTAGT

     TrpGlyAlaAspThrAlaAlaCysGlyAspIleIleAsnGlyLeuProValSerAlaArg
1081 CGTGGGGGGCAGATACCGCCGCGTGCGGTGACATCATCAACGGCTTGCCTGTTTCCGCCC
     GCACCCCCCGTCTATGGCGGCGCACGCCACTGTAGTAGTTGCCGAACGGACAAAGGCGGG

     ArgGlyArgGluIleLeuLeuGlyProAlaAspGlyMetValSerLysGlyTrpArgLeu
1141 GCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGGT
     CGTCCCCGGCCCTCTATGACGAGCCCGGTCGGCTACCTTACCAGAGGTTCCCCACCTCCA

     LeuAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
1201 TGCTGGCGCCCATCACGGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATCA
     ACGACCGCGGGTAGTGCCGCATGCGGGTCGTCTGTTCCCCGGAGGATCCCACGTATTAGT

     SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
1261 CCAGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTG
     GGTCGGATTGACCGGCCCTGTTTTTGGTTCACCTCCCACTCCAGGTCTAACACAGTTGAC

     AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
1321 CTGCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGGG
     GACGGGTTTGGAAGGACCGTTGCACGTAGTTACCCCACACGACCTGACAGATGGTGCCCC

     GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
1381 CCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTAG
     GGCCTTGCTCCTGGTAGCGCAGTGGGTTCCCAGGACAGTAGGTCTACATATGGTTACATC

     GlnAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
1441 ACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACTT
     TGGTTCTGGAACACCCGACCGGGCGAGGCGTTCCATCGGCGAGTAACTGTGGGACGTGAA

     GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg
1501 GCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCCGGC
     CGCCGAGGAGCCTGGAAATGGACCAGTGCTCCGTGCGGCTACAGTAAGGGCACGCGGCCG

     GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
1561 GGGGTGATAGCAGGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCCT
     CCCCACTATCGTCCCCGTCGGACGACAGCGGGGCCGGGTAAAGGATGAACTTTCCGAGGA

     GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
1621 CGGGGGGTCCGCTGTTGTGCCCCGCGGGGCACGCCGTGGGCATATTTAGGGCCGCGGTGT
     GCCCCCCAGGCGACAACACGGGGCGCCCCGTGCGGCACCCGTATAAATCCCGGCGCCACA

     ThrArgGlyValAlaLysAlaValAspPheIleProValGluAsnLeuGluThrThrMet
1681 GCACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCA
     CGTGGGCACCTCACCGATTCCGCCACCTGAAATAGGGACACCTCTTGGATCTCTGTTGGT

     ArgSerProValPheThrAspAsnSerSerProProValValProGlnSerPheGlnVal
1741 TGAGGTCCCCGGTGTTCACGGATAACTCCTCTCCACCAGTAGTGCCCCAGAGCTTCCAGG
     ACTCCAGGGGCCACAAGTGCCTATTGAGGAGAGGTGGTCATCACGGGGTCTCGAAGGTCC

     AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla
1801 TGGCTCACCTCCATGCTCCCACAGGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATG
     ACCGAGTGGAGGTACGAGGGTGTCCGTCGCCGTTTTCGTGGTTCCAGGGCCGACGTATAC

     AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrLeuGlyPheGly
1861 CAGCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTG
     GTCGAGTCCCGATATTCCACGATCATGAGTTGGGGAGACAACGACGTTGTGACCCGAAAC

     AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle
```

FIG. 32-2

```
1921 GTGCTTACATGTCCAAGGCTCATGGGATCGATCCTAACATCAGGACCGGGGTGAGAACAA
     CACGAATGTACAGGTTCCGAGTACCCTAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTT

      ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys
1981 TTACCACTGGCAGCCCCATCACGTACTCCACCTACGGCAAGTTCCTTGCCGACGGCGGGT
     AATGGTGACCGTCGGGGTAGTGCATGAGGTGGATGCCGTTCAAGGAACGGCTGCCGCCCA

      SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer
2041 GCTCGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACAT
     CGAGCCCCCCGCGAATACTGTATTATTAAACACTGCTCACGGTGAGGTGCCTACGGTGTA

      IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuValVal
2101 CCATCTTGGGCATCGGCACTGTCCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTG
     GGTAGAACCCGTAGCCGTGACAGGAACTGGTTCGTCTCTGACGCCCCCGCTCTGACCAAC

      LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal
2161 TGCTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCATCCCAACATCGAGGAGG
     ACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCC

      AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluValIle
2221 TTGCTCTGTCCACCACCGGAGAGATCCCTTTTTACGGCAAGGCTATCCCCCTCGAAGTAA
     AACGAGACAGGTGGTGGCCTCTCTAGGGAAAAATGCCGTTCCGATAGGGGGAGCTTCATT

      LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla
2281 TCAAGGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAAGAAGTGCGACGAACTCGCCG
     AGTTCCCCCCCTCTGTAGAGTAGAAGACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGC

      LysLeuValAlaLeuGlyIleAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal
2341 CAAAGCTGGTCGCATTGGGCATCAATGCCGTGGCCTACTACCGCGGTCTTGACGTGTCCG
     GTTTCGACCAGCGTAACCCGTAGTTACGGCACCGGATGATGGCGCCAGAACTGCACAGGC

      IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyTyrThr
2401 TCATCCCGACCAGCGGCGATGTTGTCGTCGTGGCAACCGATGCCCTCATGACCGGCTATA
     AGTAGGGCTGGTCGCCGCTACAACAGCAGCACCGTTGGCTACGGGAGTACTGGCCGATAT

      GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer
2461 CCGGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCACCCAGACAGTCGATTTCA
     GGCCGCTGAAGCTGAGCCACTATCTGACGTTATGCACACAGTGGGTCTGTCAGCTAAAGT

      LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr
2521 GCCTTGACCCTACCTTCACCATTGAGACAATCACGCTCCCCCAGGATGCTGTCTCCCGCA
     CGGAACTGGGATGGAAGTGGTAACTCTGTTAGTGCGAGGGGGTCCTACGACAGAGGGCGT

      GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly
2581 CTCAACGTCGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGG
     GAGTTGCAGCCCCGTCCTGACCGTCCCCCTTCGGTCCGTAGATGTCTAAACACCGTGGCC

      GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys
2641 GGGAGCGCCCCTCCGGCATGTTCGACTCGTCCGTCCTCTGTGAGTGCTATGACGCAGGCT
     CCCTCGCGGGGAGGCCGTACAAGCTGAGCAGGCAGGAGACACTCACGATACTGCGTCCGA

      AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr
2701 GTGCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACA
     CACGAACCATACTCGAGTGCGGGCGGCTCTGATGTCAATCCGATGCTCGCATGTACTTGT

      ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu
2761 CCCCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCC
     GGGGCCCCGAAGGGCACACGGTCCTGGTAGAACTTAAAACCCTCCCGCAGAAATGTCCGG

      ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr
2821 TCACTCATATAGATGCCCACTTTCTATCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTT
     AGTGAGTATATCTACGGGTGAAAGATAGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAA

      LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp
2881 ACCTGGTAGCGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCCTCCCCCATCGTGGG
     TGGACCATCGCATGGTTCGGTGGCACACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCC
```

FIG. 32-3

```
      GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu
2941  ACCAGATGTGGAAGTGTTTGATTCGCCTCAAGCCCACCCTCCATGGCCAACACCCCTGC
      TGGTCTACACCTTCACAAACTAAGCGGAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACG

      TyrArgLeuGlyAlaValGlnAsnGluIleThrLeuThrHisProValThrLysTyrIle
3001  TATACAGACTGGGCGCTGTTCAGAATGAAATCACCCTGACGCACCCAGTCACCAAATACA
      ATATGTCTGACCCGCGACAAGTCTTACTTTAGTGGGACTGCGTGGGTCAGTGGTTTATGT

      MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly
3061  TCATGACATGCATGTCGGCCGACCTGGAGGTCGTCACGAGCACCTGGGTGCTCGTTGGCG
      AGTACTGTACGTACAGCCGGCTGGACCTCCAGCAGTGCTCGTGGACCCACGAGCAACCGC

      ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg
3121  GCGTCCTGGCTGCTTTGGCCGCGTATTGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCA
      CGCAGGACCGACGAAACCGGCGCATAACGGACAGTTGTCCGACGCACCAGTATCACCCGT

      ValValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe
3181  GGGTCGTCTTGTCCGGGAAGCCGGCAATCATACCTGACAGGGAAGTCCTCTACCGAGAGT
      CCCAGCAGAACAGGCCCTTCGGCCGTTAGTATGGACTGTCCCTTCAGGAGATGGCTCTCA

      AspGluMetGluGluCysSerGlnHisLeuProTyrIleGluGlnGlyMetMetLeuAla
3241  TCGATGAGATGGAAGAGTGCTCTCAGCACTTACCGTACATCGAGCAAGGGATGATGCTCG
      AGCTACTCTACCTTCTCACGAGAGTCGTGAATGGCATGTAGCTCGTTCCCTACTACGAGC

      GluGlnPheLysGlnLysAlaLeuGlyLeuLeuGlnThrAlaSerArgGlnAlaGluVal
3301  CCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGG
      GGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCC

      IleAlaProAlaValGlnThrAsnTrpGlnLysLeuGluThrPheTrpAlaLysHisMet
3361  TTATCGCCCCTGCTGTCCAGACCAACTGGCAAAAACTCGAGACCTTCTGGGCGAAGCATA
      AATAGCGGGGACGACAGGTCTGGTTGACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTAT

      TrpAsnPheIleSerGlyIleGlnTyrLeuAlaGlyLeuSerThrLeuProGlyAsnPro
3421  TGTGGAACTTCATCAGTGGGATACAATACTTGGCGGGCTTGTCAACGCTGCCTGGTAACC
      ACACCTTGAAGTAGTCACCCTATGTTATGAACCGCCCGAACAGTTGCGACGGACCATTGG

      AlaIleAlaSerLeuMetAlaPheThrAlaAlaValThrSerProLeuThrThrSerGln
3481  CCGCCATTGCTTCATTGATGGCTTTTACAGCTGCTGTCACCAGCCCACTAACCACTAGCC
      GGCGGTAACGAAGTAACTACCGAAAATGTCGACGACAGTGGTCGGGTGATTGGTGATCGG

      ThrLeuLeuPheAsnIleLeuGlyGlyTrpValAlaAlaGlnLeuAlaAlaProGlyAla
3541  AAACCCTCCTCTTCAACATATTGGGGGGGTGGGTGGCTGCCCAGCTCGCCGCCCCCGGTG
      TTTGGGAGGAGAAGTTGTATAACCCCCCCACCCACCGACGGGTCGAGCGGCGGGGGCCAC

      AlaThrAlaPheValGlyAlaGlyLeuAlaGlyAlaAlaIleGlySerValGlyLeuGly
3601  CCGCTACTGCCTTTGTGGGCGCTGGCTTAGCTGGCGCCGCCATCGGCAGTGTTGGACTGG
      GGCGATGACGGAAACACCCGCGACCGAATCGACCGCGGCGGTAGCCGTCACAACCTGACC

      LysValLeuIleAspIleLeuAlaGlyTyrGlyAlaGlyValAlaGlyAlaLeuValAla
3661  GGAAGGTCCTCATAGACATCCTTGCAGGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGG
      CCTTCCAGGAGTATCTGTAGGAACGTCCCATACCGCGCCCGCACCGCCCTCGAGAACACC

      PheLysIleMetSerGlyGluValProSerThrGluAspLeuValAsnLeuLeuProAla
3721  CATTCAAGATCATGAGCGGTGAGGTCCCCTCCACGGAGGACCTGGTCAATCTACTGCCCG
      GTAAGTTCTAGTACTCGCCACTCCAGGGGAGGTGCCTCCTGGACCAGTTAGATGACGGGC

      IleLeuSerProGlyAlaLeuValValGlyValValCysAlaAlaIleLeuArgArgHis
3781  CCATCCTCTCGCCCGGAGCCCTCGTAGTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGC
      GGTAGGAGAGCGGGCCTCGGGAGCATCAGCCGCACCAGACACGTCGTTATGACGCGGCCG

      ValGlyProGlyGluGlyAlaValGlnTrpMetAsnArgLeuIleAlaPheAlaSerArg
3841  ACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGGATGAACCGGCTGATAGCCTTCGCCTCCC
      TGCAACCGGGCCCGCTCCCCCGTCACGTCACCTACTTGGCCGACTATCGGAAGCGGAGGG

      GlyAsnHisValSerProThrHisTyrValProGluSerAspAlaAlaAlaArgValThr
```

FIG. 32-4

```
3901 GGGGGAACCATGTTTCCCCCACGCACTACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCA
     CCCCCTTGGTACAAAGGGGGTGCGTGATGCACGGCCTCTCGCTACGTCGACGGGCGCAGT

      AlaIleLeuSerSerLeuThrValThrGlnLeuLeuArgArgLeuHisGlnTrpIleSer
3961 CTGCCATACTCAGCAGCCTCACTGTAACCCAGCTCCTGAGGCGACTGCACCAGTGGATAA
     GACGGTATGAGTCGTCGGAGTGACATTGGGTCGAGGACTCCGCTGACGTGGTCACCTATT

      SerGluCysThrThrProCysSerGlySerTrpLeuArgAspIleTrpAspTrpIleCys
4021 GCTCGGAGTGTACCACTCCATGCTCCGGTTCCTGGCTAAGGGACATCTGGGACTGGATAT
     CGAGCCTCACATGGTGAGGTACGAGGCCAAGGACCGATTCCCTGTAGACCCTGACCTATA

      GluValLeuSerAspPheLysThrTrpLeuLysAlaLysLeuMetProGlnLeuProGly
4081 GCGAGGTGTTGAGCGACTTTAAGACCTGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTG
     CGCTCCACAACTCGCTGAAATTCTGGACCGATTTTCGATTCGAGTACGGTGTCGACGGAC

      IleProPheValSerCysGlnArgGlyTyrLysGlyValTrpArgValAspGlyIleMet
4141 GGATCCCCTTTGTGTCCTGCCAGCGCGGGTATAAGGGGGTCTGGCGAGTGGACGGCATCA
     CCTAGGGGAAACACAGGACGGTCGCGCCCATATTCCCCCAGACCGCTCACCTGCCGTAGT

      HisThrArgCysHisCysGlyAlaGluIleThrGlyHisValLysAsnGlyThrMetArg
4201 TGCACACTCGCTGCCACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACGATGA
     ACGTGTGAGCGACGGTGACACCTCGACTCTAGTGACCTGTACAGTTTTTGCCCTGCTACT

      IleValGlyProArgThrCysArgAsnMetTrpSerGlyThrPheProIleAsnAlaTyr
4261 GGATCGTCGGTCCTAGGACCTGCAGGAACATGTGGAGTGGGACCTTCCCCATTAATGCCT
     CCTAGCAGCCAGGATCCTGGACGTCCTTGTACACCTCACCCTGGAAGGGGTAATTACGGA

      ThrThrGlyProCysThrProLeuProAlaProAsnTyrThrPheAlaLeuTrpArgVal
4321 ACACCACGGGCCCCTGTACCCCCCTTCCTGCGCCGAACTACACGTTCGCGCTATGGAGGG
     TGTGGTGCCCGGGGACATGGGGGGAAGGACGCGGCTTGATGTGCAAGCGCGATACCTCCC

      SerAlaGluGluTyrValGluIleArgGlnValGlyAspPheHisTyrValThrGlyMet
4381 TGTCTGCAGAGGAATATGTGGAGATAAGGCAGGTGGGGGACTTCCACTACGTGACGGGTA
     ACAGACGTCTCCTTATACACCTCTATTCCGTCCACCCCCTGAAGGTGATGCACTGCCCAT

      ThrThrAspAsnLeuLysCysProCysGlnValProSerProGluPhePheThrGluLeu
4441 TGACTACTGACAATCTCAAATGCCCGTGCCAGGTCCCATCGCCCGAATTTTTCACAGAAT
     ACTGATGACTGTTAGAGTTTACGGGCACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTA

      AspGlyValArgLeuHisArgPheAlaProProCysLysProLeuLeuArgGluGluVal
4501 TGGACGGGGTGCGCCTACATAGGTTTGCGCCCCCCTGCAAGCCCTTGCTGCGGGAGGAGG
     ACCTGCCCCACGCGGATGTATCCAAACGCGGGGGGACGTTCGGGAACGACGCCCTCCTCC

      SerPheArgValGlyLeuHisGluTyrProValGlySerGlnLeuProCysGluProGlu
4561 TATCATTCAGAGTAGGACTCCACGAATACCCGGTAGGGTCGCAATTACCTTGCGAGCCCG
     ATAGTAAGTCTCATCCTGAGGTGCTTATGGGCCATCCCAGCGTTAATGGAACGCTCGGGC

      ProAspValAlaValLeuThrSerMetLeuThrAspProSerHisIleThrAlaGluAla
4621 AACCGGACGTGGCCGTGTTGACGTCCATGCTCACTGATCCCTCCCATATAACAGCAGAGG
     TTGGCCTGCACCGGCACAACTGCAGGTACGAGTGACTAGGGAGGGTATATTGTCGTCTCC

      AlaGlyArgArgLeuAlaArgGlySerProProSerValAlaSerSerSerAlaSerGln
4681 CGGCCGGGCGAAGGTTGGCGAGGGGATCACCCCCCTCTGTGGCCAGCTCCTCGGCTAGCC
     GCCGGCCCGCTTCCAACCGCTCCCCTAGTGGGGGGAGACACCGGTCGAGGAGCCGATCGG

      LeuSerAlaProSerLeuLysAlaThrCysThrAlaAsnHisAspSerProAspAlaGlu
4741 AGCTATCCGCTCCATCTCTCAAGGCAACTTGCACCGCTAACCATGACTCCCCTGATGCTG
     TCGATAGGCGAGGTAGAGAGTTCCGTTGAACGTGGCGATTGGTACTGAGGGGACTACGAC

      LeuIleGluAlaAsnLeuLeuTrpArgGlnGluMetGlyGlyAsnIleThrArgValGlu
4801 AGCTCATAGAGGCCAACCTCCTATGGAGGCAGGAGATGGGCGGCAACATCACCAGGGTTG
     TCGAGTATCTCCGGTTGGAGGATACCTCCGTCCTCTACCCGCCGTTGTAGTGGTCCCAAC

      SerGluAsnLysValValIleLeuAspSerPheAspProLeuValAlaGluGluAspGlu
4861 AGTCAGAAAACAAAGTGGTGATTCTGGACTCCTTCGATCCGCTTGTGGCGGAGGAGGACG
     TCAGTCTTTTGTTTCACCACTAAGACCTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGC
```

FIG. 32-5

```
      ArgGluIleSerValProAlaGluIleLeuArgLysSerArgArgPheAlaGlnAlaLeu
4921  AGCGGGAGATCTCCGTACCCGCAGAAATCCTGCGGAAGTCTCGGAGATTCGCCCAGGCCC
      TCGCCCTCTAGAGGCATGGGCGTCTTTAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGG

      ProValTrpAlaArgProAspTyrAsnProProLeuValGluThrTrpLysLysProAsp
4981  TGCCCGTTTGGGCGCGGCCGGACTATAACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCG
      ACGGGCAAACCCGCGCCGGCCTGATATTGGGGGGCGATCACCTCTGCACCTTTTTCGGGC

      TyrGluProProValValHisGlyCysProLeuProProProLysSerProProValPro
5041  ACTACGAACCACCTGTGGTCCATGGCTGTCCGCTTCCACCTCCAAAGTCCCCTCCTGTGC
      TGATGCTTGGTGGACACCAGGTACCGACAGGCGAAGGTGGAGGTTTCAGGGGAGGACACG

      ProProArgLysLysArgThrValValLeuThrGluSerThrLeuSerThrAlaLeuAla
5101  CTCCGCCTCGGAAGAAGCGGACGGTGGTCCTCACTGAATCAACCCTATCTACTGCCTTGG
      GAGGCGGAGCCTTCTTCGCCTGCCACCAGGAGTGACTTAGTTGGGATAGATGACGGAACC

      GluLeuAlaThrArgSerPheGlySerSerSerThrSerGlyIleThrGlyAspAsnThr
5161  CCGAGCTCGCCACCAGAAGCTTTGGCAGCTCCTCAACTTCCGGCATTACGGGCGACAATA
      GGCTCGAGCGGTGGTCTTCGAAACCGTCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTAT

      ThrThrSerSerGluProAlaProSerGlyCysProProAspSerAspAlaGluSerTyr
5221  CGACAACATCCTCTGAGCCCGCCCCTTCTGGCTGCCCCCCCGACTCCGACGCTGAGTCCT
      GCTGTTGTAGGAGACTCGGGCGGGGAAGACCGACGGGGGGGCTGAGGCTGCGACTCAGGA

      SerSerMetProProLeuGluGlyGluProGlyAspProAspLeuSerAspGlySerTrp
5281  ATTCCTCCATGCCCCCCCTGGAGGGGGAGCCTGGGGATCCGGATCTTAGCGACGGGTCAT
      TAAGGAGGTACGGGGGGGACCTCCCCCTCGGACCCCTAGGCCTAGAATCGCTGCCCAGTA

      SerThrValSerSerGluAlaAsnAlaGluAspValValCysCysSerMetSerTyrSer
5341  GGTCAACGGTCAGTAGTGAGGCCAACGCGGAGGATGTCGTGTGCTGCTCAATGTCTTACT
      CCAGTTGCCAGTCATCACTCCGGTTGCGCCTCCTACAGCACACGACGAGTTACAGAATGA

      TrpThrGlyAlaLeuValThrProCysAlaAlaGluGluGlnLysLeuProIleAsnAla
5401  CTTGGACAGGCGCACTCGTCACCCCGTGCGCCGCGGAAGAACAGAAACTGCCCATCAATG
      GAACCTGTCCGCGTGAGCAGTGGGGCACGCGGCGCCTTCTTGTCTTTGACGGGTAGTTAC

      LeuSerAsnSerLeuLeuArgHisHisAsnLeuValTyrSerThrThrSerArgSerAla
5461  CACTAAGCAACTCGTTGCTACGTCACCACAATTTGGTGTATTCCACCACCTCACGCAGTG
      GTGATTCGTTGAGCAACGATGCAGTGGTGTTAAACCACATAAGGTGGTGGAGTGCGTCAC

      CysGlnArgGlnLysLysValThrPheAspArgLeuGlnValLeuAspSerHisTyrGln
5521  CTTGCCAAAGGCAGAAGAAAGTCACATTTGACAGACTGCAAGTTCTGGACAGCCATTACC
      GAACGGTTTCCGTCTTCTTTCAGTGTAAACTGTCTGACGTTCAAGACCTGTCGGTAATGG

      AspValLeuLysGluValLysAlaAlaAlaSerLysValLysAlaAsnLeuLeuSerVal
5581  AGGACGTACTCAAGGAGGTTAAAGCAGCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCG
      TCCTGCATGAGTTCCTCCAATTTCGTCGCCGCAGTTTTCACTTCCGATTGAACGATAGGC

      GluGluAlaCysSerLeuThrProProHisSerAlaLysSerLysPheGlyTyrGlyAla
5641  TAGAGGAAGCTTGCAGCCTGACGCCCCCACACTCAGCCAAATCCAAGTTTGGTTATGGGG
      ATCTCCTTCGAACGTCGGACTGCGGGGGTGTGAGTCGGTTTAGGTTCAAACCAATACCCC

      LysAspValArgCysHisAlaArgLysAlaValThrHisIleAsnSerValTrpLysAsp
5701  CAAAAGACGTCCGTTGCCATGCCAGAAAGGCCGTAACCCACATCAACTCCGTGTGGAAAG
      GTTTTCTGCAGGCAACGGTACGGTCTTTCCGGCATTGGGTGTAGTTGAGGCACACCTTTC

      LeuLeuGluAspAsnValThrProIleAspThrThrIleMetAlaLysAsnGluValPhe
5761  ACCTTCTGGAAGACAATGTAACACCAATAGACACTACCATCATGGCTAAGAACGAGGTTT
      TGGAAGACCTTCTGTTACATTGTGGTTATCTGTGATGGTAGTACCGATTCTTGCTCCAAA

      CysValGlnProGluLysGlyGlyArgLysProAlaArgLeuIleValPheProAspLeu
5821  TCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAGCCAGCTCGTCTCATCGTGTTCCCCGATC
      AGACGCAAGTCGGACTCTTCCCCCCAGCATTCGGTCGAGCAGAGTAGCACAAGGGGCTAG

      GlyValArgValCysGluLysMetAlaLeuTyrAspValValThrLysLeuProLeuAla
```

FIG. 32-6

```
5881  TGGGCGTGCGCGTGTGCGAAAAGATGGCTTTGTACGACGTGGTTACAAAGCTCCCCTTGG
      ACCCGCACGCGCACACGCTTTTCTACCGAAACATGCTGCACCAATGTTTCGAGGGGAACC

       ValMetGlySerSerTyrGlyPheGlnTyrSerProGlyGlnArgValGluPheLeuVal
5941  CCGTGATGGGAAGCTCCTACGGATTCCAATACTCACCAGGACAGCGGGTTGAATTCCTCG
      GGCACTACCCTTCGAGGATGCCTAAGGTTATGAGTGGTCCTGTCGCCCAACTTAAGGAGC

       GlnAlaTrpLysSerLysLysThrProMetGlyPheSerTyrAspThrArgCysPheAsp
6001  TGCAAGCGTGGAAGTCCAAGAAAACCCCAATGGGGTTCTCGTATGATACCCGCTGCTTTG
      ACGTTCGCACCTTCAGGTTCTTTTGGGGTTACCCCAAGAGCATACTATGGGCGACGAAAC

       SerThrValThrGluSerAspIleArgThrGluGluAlaIleTyrGlnCysCysAspLeu
6061  ACTCCACAGTCACTGAGAGCGACATCCGTACGGAGGAGGCAATCTACCAATGTTGTGACC
      TGAGGTGTCAGTGACTCTCGCTGTAGGCATGCCTCCTCCGTTAGATGGTTACAACACTGG

       AspProGlnAlaArgValAlaIleLysSerLeuThrGluArgLeuTyrValGlyGlyPro
6121  TCGACCCCCAAGCCCGCGTGGCCATCAAGTCCCTCACCGAGAGGCTTTATGTTGGGGGCC
      AGCTGGGGGTTCGGGCGCACCGGTAGTTCAGGGAGTGGCTCTCCGAAATACAACCCCCGG

       LeuThrAsnSerArgGlyGluAsnCysGlyTyrArgArgCysArgAlaSerGlyValLeu
6181  CTCTTACCAATTCAAGGGGGGAGAACTGCGGCTATCGCAGGTGCCGCGCGAGCGGCGTAC
      GAGAATGGTTAAGTTCCCCCCTCTTGACGCCGATAGCGTCCACGGCGCGCTCGCCGCATG

       ThrThrSerCysGlyAsnThrLeuThrCysTyrIleLysAlaArgAlaAlaCysArgAla
6241  TGACAACTAGCTGTGGTAACACCCTCACTTGCTACATCAAGGCCCGGGCAGCCTGTCGAG
      ACTGTTGATCGACACCATTGTGGGAGTGAACGATGTAGTTCCGGGCCCGTCGGACAGCTC

       AlaGlyLeuGlnAspCysThrMetLeuValCysGlyAspAspLeuValValIleCysGlu
6301  CCGCAGGGCTCCAGGACTGCACCATGCTCGTGTGTGGCGACGACTTAGTCGTTATCTGTG
      GGCGTCCCGAGGTCCTGACGTGGTACGAGCACACACCGCTGCTGAATCAGCAATAGACAC

       SerAlaGlyValGlnGluAspAlaAlaSerLeuArgAlaPheThrGluAlaMetThrArg
6361  AAAGCGCGGGGGTCCAGGAGGACGCGGCGAGCCTGAGAGCCTTCACGGAGGCTATGACCA
      TTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCGGACTCTCGGAAGTGCCTCCGATACTGGT

       TyrSerAlaProProGlyAspProProGlnProGluTyrAspLeuGluLeuIleThrSer
6421  GGTACTCCGCCCCCCCTGGGGACCCCCCACAACCAGAATACGACTTGGAGCTCATAACAT
      CCATGAGGCGGGGGGGACCCCTGGGGGGTGTTGGTCTTATGCTGAACCTCGAGTATTGTA

       CysSerSerAsnValSerValAlaHisAspGlyAlaGlyLysArgValTyrTyrLeuThr
6481  CATGCTCCTCCAACGTGTCAGTCGCCCACGACGGCGCTGGAAAGAGGGTCTACTACCTCA
      GTACGAGGAGGTTGCACAGTCAGCGGGTGCTGCCGCGACCTTTCTCCCAGATGATGGAGT

       ArgAspProThrThrProLeuAlaArgAlaAlaTrpGluThrAlaArgHisThrProVal
6541  CCCGTGACCCTACAACCCCCCTCGCGAGAGCTGCGTGGGAGACAGCAAGACACACTCCAG
      GGGCACTGGGATGTTGGGGGGAGCGCTCTCGACGCACCCTCTGTCGTTCTGTGTGAGGTC

       AsnSerTrpLeuGlyAsnIleIleMetPheAlaProThrLeuTrpAlaArgMetIleLeu
6601  TCAATTCCTGGCTAGGCAACATAATCATGTTTGCCCCCACACTGTGGGCGAGGATGATAC
      AGTTAAGGACCGATCCGTTGTATTAGTACAAACGGGGGTGTGACACCCGCTCCTACTATG

       MetThrHisPhePheSerValLeuIleAlaArgAspGlnLeuGluGlnAlaLeuAspCys
6661  TGATGACCCATTTCTTTAGCGTCCTTATAGCCAGGGACCAGCTTGAACAGGCCCTCGATT
      ACTACTGGGTAAAGAAATCGCAGGAATATCGGTCCCTGGTCGAACTTGTCCGGGAGCTAA

       GluIleTyrGlyAlaCysTyrSerIleGluProLeuAspLeuProProIleIleGlnArg
6721  GCGAGATCTACGGGGCCTGCTACTCCATAGAACCACTTGATCTACCTCCAATCATTCAAA
      CGCTCTAGATGCCCCGGACGATGAGGTATCTTGGTGAACTAGATGGAGGTTAGTAAGTTT

       Leu
6781  GACTC
      CTGAG
```

FIG. 32-7

**FIG. 47-1** COMBINED ORF OF DNAs K9-1 through 15e

```
     GlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGlyTrpGly
  1  CAGGCTGTCCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTTGACCAGGGCTGGG
     GTCCGACAGGACTCTCCGATCGGTCGACGGCTGGGGAATGGCTAAAACTGGTCCCGACCC

     ProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHisTyrPro
 61  GCCCTATCAGTTATGCCAACGGAAGCGGCCCCGACCAGCGCCCCTACTGCTGGCACTACC
     CGGGATAGTCAATACGGTTGCCTTCGCCGGGGCTGGTCGCGGGGATGACGACCGTGATGG

     ProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCysPheThr
121  CCCCAAAACCTTGCGGTATTGTGCCCGCGAAGAGTGTGTGTGGTCCGGTATATTGCTTCA
     GGGGTTTTGGAACGCCATAACACGGGCGCTTCTCACACACACCAGGCCATATAACGAAGT

     ProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSerTrpGly
181  CTCCCAGCCCCGTGGTGGTGGGAACGACCGACAGGTCGGGCGCGCCCACCTACAGCTGGG
     GAGGGTCGGGGCACCACCACCCTTGCTGGCTGTCCAGCCCGCGCGGGTGGATGTCGACCC

     GluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsnTrpPhe
241  GTGAAAATGATACGGACGTCTTCGTCCTTAACAATACCAGGCCACCGCTGGGCAATTGGT
     CACTTTTACTATGCCTGCAGAAGCAGGAATTGTTATGGTCCGGTGGCGACCCGTTAACCA

     GlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProProCysVal
301  TCGGTTGTACCTGGATGAACTCAACTGGATTCACCAAAGTGTGCGGAGCGCCTCCTTGTG
     AGCCAACATGGACCTACTTGAGTTGACCTAAGTGGTTTCACACGCCTCGCGGAGGAACAC

     IleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLysHisPro
361  TCATCGGAGGGGCGGGCAACAACACCCTGCACTGCCCCACTGATTGCTTCCGCAAGCATC
     AGTAGCCTCCCCGCCCGTTGTTGTGGGACGTGACGGGGTGACTAACGAAGGCGTTCGTAG

     AspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeuValAsp
421  CGGACGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTGCCTGGTCG
     GCCTGCGGTGTATGAGAGCCACGCCGAGGCCAGGGACCTAGTGTGGGTCCACGGACCAGC

     TyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLysIleArg
481  ACTACCCGTATAGGCTTTGGCATTATCCTTGTACCATCAACTACACCATATTTAAAATCA
     TGATGGGCATATCCGAAACCGTAATAGGAACATGGTAGTTGATGTGGTATAAATTTTAGT

     MetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArgGlyGlu
541  GGATGTACGTGGGAGGGGTCGAACACAGGCTGGAAGCTGCCTGCAACTGGACGCGGGGCG
     CCTACATGCACCCTCCCCAGCTTGTGTCCGACCTTCGACGGACGTTGACCTGCGCCCCGC

     ArgCysAspLeuGluAspArgAspArgSerGluLeuSerProLeuLeuLeuThrThrThr
601  AACGTTGCGATCTGGAAGACAGGGACAGGTCCGAGCTCAGCCCGTTACTGCTGACCACTA
     TTGCAACGCTAGACCTTCTGTCCCTGTCCAGGCTCGAGTCGGGCAATGACGACTGGTGAT

     GlnTrpGlnValLeuProCysSerPheThrThrLeuProAlaLeuSerThrGlyLeuIle
661  CACAGTGGCAGGTCCTCCCGTGTTCCTTCACAACCCTACCAGCCTTGTCCACCGGCCTCA
     GTGTCACCGTCCAGGAGGGCACAAGGAAGTGTTGGGATGGTCGGAACAGGTGGCCGGAGT

     HisLeuHisGlnAsnIleValAspValGlnTyrLeuTyrGlyValGlySerSerIleAla
721  TCCACCTCCACCAGAACATTGTGGACGTGCAGTACTTGTACGGGGTGGGGTCAAGCATCG
     AGGTGGAGGTGGTCTTGTAACACCTGCACGTCATGAACATGCCCCACCCCAGTTCGTAGC

     SerTrpAlaIleLysTrpGluTyrValValLeuLeuPheLeuLeuLeuAlaAspAlaArg
781  CGTCCTGGGCCATTAAGTGGGAGTACGTCGTTCTCCTGTTCCTTCTGCTTGCAGACGCGC
     GCAGGACCCGGTAATTCACCCTCATGCAGCAAGAGGACAAGGAAGACGAACGTCTGCGCG

     ValCysSerCysLeuTrpMetMetLeuLeuIleSerGlnAlaGluAlaAlaLeuGluAsn
841  GCGTCTGCTCCTGCTTGTGGATGATGCTACTCATATCCCAAGCGGAGGCGGCTTTGGAGA
     CGCAGACGAGGACGAACACCTACTACGATGAGTATAGGGTTCGCCTCCGCCGAAACCTCT

     LeuValIleLeuAsnAlaAlaSerLeuAlaGlyThrHisGlyLeuValSerPheLeuVal
901  ACCTCGTAATACTTAATGCAGCATCCCTGGCCGGGACGCACGGTCTTGTATCCTTCCTCG
     TGGAGCATTATGAATTACGTCGTAGGGACCGGCCCTGCGTGCCAGAACATAGGAAGGAGC
```

PhePheCysPheAlaTrpTyrLeuLysGlyLysTrpValProGlyAlaValTyrThrPhe
961 TGTTCTTCTGCTTTGCATGGTATTTGAAGGGTAAGTGGGTGCCCGGAGCGGTCTACACCT
ACAAGAAGACGAAACGTACCATAAACTTCCCATTCACCCACGGGCCTCGCCAGATGTGGA

TyrGlyMetTrpProLeuLeuLeuLeuLeuLeuAlaLeuProGlnArgAlaTyrAlaLeu
1021 TCTACGGGATGTGGCCTCTCCTCCTGCTCCTGTTGGCGTTGCCCCAGCGGGCGTACGCGC
AGATGCCCTACACCGGAGAGGAGGACGAGGACAACCGCAACGGGGTCGCCCGCATGCGCG

AspThrGluValAlaAlaSerCysGlyGlyValValLeuValGlyLeuMetAlaLeuThr
1081 TGGACACGGAGGTGGCCGCGTCGTGTGGCGGTGTTGTTCTCGTCGGGTTGATGGCGCTGA
ACCTGTGCCTCCACCGGCGCAGCACACCGCCACAACAAGAGCAGCCCAACTACCGCGACT

LeuSerProTyrTyrLysArgTyrIleSerTrpCysLeuTrpTrpLeuGlnTyrPheLeu
1141 CTCTGTCACCATATTACAAGCGCTATATCAGCTGGTGCTTGTGGTGGCTTCAGTATTTTC
GAGACAGTGGTATAATGTTCGCGATATAGTCGACCACGAACACCACCGAAGTCATAAAAG

ThrArgValGluAlaGlnLeuHisValTrpIleProProLeuAsnValArgGlyGlyArg
1201 TGACCAGAGTGGAAGCGCAACTGCACGTGTGGATTCCCCCCCTCAACGTCCGAGGGGGGC
ACTGGTCTCACCTTCGCGTTGACGTGCACACCTAAGGGGGGGAGTTGCAGGCTCCCCCCG

AspAlaValIleLeuLeuMetCysAlaValHisProThrLeuValPheAspIleThrLys
1261 GCGACGCCGTCATCTTACTCATGTGTGCTGTACACCCGACTCTGGTATTTGACATCACCA
CGCTGCGGCAGTAGAATGAGTACACACGACATGTGGGCTGAGACCATAAACTGTAGTGGT

LeuLeuLeuAlaValPheGlyProLeuTrpIleLeuGlnAlaSerLeuLeuLysValPro
1321 AATTGCTGCTGGCCGTCTTCGGACCCCTTTGGATTCTTCAAGCCAGTTTGCTTAAAGTAC
TTAACGACGACCGGCAGAAGCCTGGGGAAACCTAAGAAGTTCGGTCAAACGAATTTCATG

TyrPheValArgValGlnGlyLeuLeuArgPheCysAlaLeuAlaArgLysMetIleGly
1381 CCTACTTTGTGCGCGTCCAAGGCCTTCTCCGGTTCTGCGCGTTAGCGCGGAAGATGATCG
GGATGAAACACGCGCAGGTTCCGGAAGAGGCCAAGACGCGCAATCGCGCCTTCTACTAGC

GlyHisTyrValGlnMetValIleIleLysLeuGlyAlaLeuThrGlyThrTyrValTyr
1441 GAGGCCATTACGTGCAAATGGTCATCATTAAGTTAGGGGCGCTTACTGGCACCTATGTTT
CTCCGGTAATGCACGTTTACCAGTAGTAATTCAATCCCCGCGAATGACCGTGGATACAAA

AsnHisLeuThrProLeuArgAspTrpAlaHisAsnGlyLeuArgAspLeuAlaValAla
1501 ATAACCATCTCACTCCTCTTCGGGACTGGGCGCACAACGGCTTGCGAGATCTGGCCGTGG
TATTGGTAGAGTGAGGAGAAGCCCTGACCCGCGTGTTGCCGAACGCTCTAGACCGGCACC

ValGluProValValPheSerGlnMetGluThrLysLeuIleThrTrpGlyAlaAspThr
1561 CTGTAGAGCCAGTCGTCTTCTCCCAAATGGAGACCAAGCTCATCACGTGGGGGGCAGATA
GACATCTCGGTCAGCAGAAGAGGGTTTACCTCTGGTTCGAGTAGTGCACCCCCCGTCTAT

AlaAlaCysGlyAspIleIleAsnGlyLeuProValSerAlaArgArgGlyArgGluIle
1621 CCGCCGCGTGCGGTGACATCATCAACGGCTTGCCTGTTTCCGCCCGCAGGGGCCGGGAGA
GGCGGCGCACGCCACTGTAGTAGTTGCCGAACGGACAAAGGCGGGCGTCCCCGGCCCTCT

LeuLeuGlyProAlaAspGlyMetValSerLysGlyTrpArgLeuLeuAlaProIleThr
1681 TACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGGTTGCTGGCGCCCATCA
ATGACGAGCCCGGTCGGCTACCTTACCAGAGGTTCCCCACCTCCAACGACCGCGGGTAGT

AlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThrSerLeuThrGlyArg
1741 CGGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATCACCAGCCTAACTGGCC
GCCGCATGCGGGTCGTCTGTTCCCCGGAGGATCCCACGTATTAGTGGTCGGATTGACCGG

AspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeu
1801 GGGACAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCTGCCCAAACCTTCC
CCCTGTTTTTGGTTCACCTCCCACTCCAGGTCTAACACAGTTGACGACGGGTTTGGAAGG

AlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIle
1861 TGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGGGCCGGAACGAGGACCA
ACCGTTGCACGTAGTTACCCCACACGACCTGACAGATGGTGCCCCGGCCTTGCTCCTGGT

AlaSerProLysGlyProValIleGlnMetTyrThrAsnValAspGlnAspLeuValGly

FIG. 47-2

```
1921  TCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTAGACCAAGACCTTGTGG
      AGCGCAGTGGGTTCCCAGGACAGTAGGTCTACATATGGTTACATCTGGTTCTGGAACACC

        TrpProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu
1981  GCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACTTGCGGCTCCTCGGACC
      CGACCGGGCGAGGCGTTCCATCGGCGAGTAACTGTGGGACGTGAACGCCGAGGAGCCTGG

        TyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGly
2041  TTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCCGGCGGGGTGATAGCAGGG
      AAATGGACCAGTGCTCCGTGCGGCTACAGTAAGGGCACGCGGCCGCCCCACTATCGTCCC

        SerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeu
2101  GCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCCTCGGGGGGTCCGCTGT
      CGTCGGACGACAGCGGGGCCGGGTAAAGGATGAACTTTCCGAGGAGCCCCCCAGGCGACA

        CysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAla
2161  TGTGCCCCGCGGGGCACGCCGTGGGCATATTTAGGGCCGCGGTGTGCACCCGTGGAGTGG
      ACACGGGGCGCCCCGTGCGGCACCCGTATAAATCCCGGCGCCACACGTGGGCACCTCACC

        LysAlaValAspPheIleProValGluAsnLeuGluThrThrMetArgSerProValPhe
2221  CTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCATGAGGTCCCCGGTGT
      GATTCCGCCACCTGAAATAGGGACACCTCTTGGATCTCTGTTGGTACTCCAGGGGCCACA

        ThrAspAsnSerSerProProValValProGlnSerPheGlnValAlaHisLeuHisAla
2281  TCACGGATAACTCCTCTCCACCAGTAGTGCCCCAGAGCTTCCAGGTGGCTCACCTCCATG
      AGTGCCTATTGAGGAGAGGTGGTCATCACGGGGTCTCGAAGGTCCACCGAGTGGAGGTAC

        ProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAlaAlaGlnGlyTyrLys
2341  CTCCCACAGGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCAGCTCAGGGCTATA
      GAGGGTGTCCGTCGCCGTTTTCGTGGTTCCAGGGCCGACGTATACGTCGAGTCCCGATAT

        ValLeuValLeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLys
2401  AGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCA
      TCCACGATCATGAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGT

        AlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerPro
2461  AGGCTCATGGGATCGATCCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCC
      TCCGAGTACCCTAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGG

        IleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyr
2521  CCATCACGTACTCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTT
      GGTAGTGCATGAGGTGGATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCCGCGAA

        AspIleIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGly
2581  ATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATCG
      TACTGTATTATTAAACACTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAGC

        ThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThr
2641  GCACTGTCCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCA
      CGTGACAGGAACTGGTTCGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGT

        ProProGlySerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThr
2701  CCCCTCCGGGCTCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCA
      GGGGAGGCCCGAGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGT

        GlyGluIleProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHis
2761  CCGGAGAGATCCCTTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGAC
      GGCCTCTCTAGGGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCTCTG

        LeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeu
2821  ATCTCATCTTCTGTCATTCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCAT
      TAGAGTAGAAGACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTA

        GlyIleAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGly
2881  TGGGCATCAATGCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCG
      ACCCGTAGTTACGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGC
```

FIG. 47-3

```
      AspValValValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSer
2941  GCGATGTTGTCGTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACT
      CGCTACAACAGCAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGA

      ValIleAspCysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPhe
3001  CGGTGATAGACTGCAATACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCT
      GCCACTATCTGACGTTATGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGA

      ThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArg
3061  TCACCATTGAGACAATCACGCTCCCCCAGGATGCTGTCTCCCGCACTCAACGTCGGGGCA
      AGTGGTAACTCTGTTAGTGCGAGGGGGTCCTACGACAGAGGGCGTGAGTTGCAGCCCCGT

      ThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGly
3121  GGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCG
      CCTGACCGTCCCCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGC

      MetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeu
3181  GCATGTTCGACTCGTCCGTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGC
      CGTACAAGCTGAGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCG

      ThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProVal
3241  TCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCG
      AGTGCGGGCGGCTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGC

      CysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAla
3301  TGTGCCAGGACCATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATG
      ACACGGTCCTGGTAGAACTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTAC

      HisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGln
3361  CCCACTTTCTATCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACC
      GGGTGAAAGATAGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGG

      AlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCys
3421  AAGCCACCGTGTGCGCTAGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGT
      TTCGGTGGCACACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCA

      LeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAla
3481  GTTTGATTCGCCTCAAGCCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCG
      CAAACTAAGCGGAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGC

      ValGlnAsnGluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSer
3541  CTGTTCAGAATGAAATCACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGT
      GACAAGTCTTACTTTAGTGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACGTACA

      AlaAspLeuGluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeu
3601  CGGCCGACCTGGAGGTCGTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTT
      GCCGGCTGGACCTCCAGCAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAA

      AlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGly
3661  TGGCCGCGTATTGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCG
      ACCGGCGCATAACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGC

      LysProAlaIleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGlu
3721  GGAAGCCGGCAATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAG
      CCTTCGGCCGTTAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTC

      CysSerGlnHisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGln
3781  AGTGCTCTCAGCACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGC
      TCACGAGAGTCGTGAATGGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCG

      LysAlaLeuGlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaVal
3841  AGAAGGCCCTCGGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTG
      TCTTCCGGGAGCCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGAC

      GlnThrAsnTrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSer
```

FIG. 47-4

3901 TCCAGACCAACTGGCAAAAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTTCATCA
AGGTCTGGT1GACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGT

GlyIleGlnTyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeu
3961 GTGGGATACAATACTTGGCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCAT
CACCCTATGTTATGAACCGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTA

MetAlaPheThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsn
4021 TGATGGCTTTTACAGCTGCTGTCACCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCA
ACTACCGAAAATGTCGACGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGT

IleLeuGlyGlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheVal
4081 ACATATTGGGGGGGTGGGTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTG
TGTATAACCCCCCCACCCACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAAC

GlyAlaGlyLeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAsp
4141 TGGGCGCTGGCTTAGCTGGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAG
ACCCGCGACCGAATCGACCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATC

IleLeuAlaGlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSer
4201 ACATCCTTGCAGGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGA
TGTAGGAACGTCCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACT

GlyGluValProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGly
4261 GCGGTGAGGTCCCCTCCACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCG
CGCCACTCCAGGGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGC

AlaLeuValValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGlu
4321 GAGCCCTCGTAGTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCG
CTCGGGAGCATCAGCCGCACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGC

GlyAlaValGlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSer
4381 AGGGGGCAGTGCAGTGGATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTT
TCCCCCGTCACGTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAA

ProThrHisTyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSer
4441 CCCCCACGCACTACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCA
GGGGGTGCGTGATGCACGGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGT

LeuThrValThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThr
4501 GCCTCACTGTAACCCAGCTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCA
CGGAGTGACATTGGGTCGAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGT

ProCysSerGlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAsp
4561 CTCCATGCTCCGGTTCCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCG
GAGGTACGAGGCCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGC

PheLysThrTrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSer
4621 ACTTTAAGACCTGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGT
TGAAATTCTGGACCGATTTTCGATTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACA

CysGlnArgGlyTyrLysGlyValTrpArgValAspGlyIleMetHisThrArgCysHis
4681 CCTGCCAGCGCGGGTATAAGGGGGTCTGGCGAGTGGACGGCATCATGCACACTCGCTGCC
GGACGGTCGCGCCCATATTCCCCCAGACCGCTCACCTGCCGTAGTACGTGTGAGCGACGG

CysGlyAlaGluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArg
4741 ACTGTGGAGCTGAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTA
TGACACCTCGACTCTAGTGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGAT

ThrCysArgAsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCys
4801 GGACCTGCAGGAACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCT
CCTGGACGTCCTTGTACACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGA

ThrProLeuProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyr
4861 GTACCCCCCTTCCTGCGCCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAAT
CATGGGGGGAAGGACGCGGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTA

FIG. 47-5

```
     ValGluIleArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeu
4921 ATGTGGAGATAAGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATC
     TACACCTCTATTCCGTCCACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAG

     LysCysProCysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeu
4981 TCAAATGCCCGTGCCAGGTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCC
     AGTTTACGGGCACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGG

     HisArgPheAlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGly
5041 TACATAGGTTTGCGCCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAG
     ATGTATCCAAACGCGGGGGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATC

     LeuHisGluTyrProValGlySerGlnLeuProCysGluProGluProAspValAlaVal
5101 GACTCCACGAATACCCGGTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCG
     CTGAGGTGCTTATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGC

     LeuThrSerMetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeu
5161 TGTTGACGTCCATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGT
     ACAACTGCAGGTACGAGTGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCA

     AlaArgGlySerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSer
5221 TGGCGAGGGGATCACCCCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCAT
     ACCGCTCCCCTAGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTA

     LeuLysAlaThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsn
5281 CTCTCAAGGCAACTTGCACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCA
     GAGAGTTCCGTTGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGT

     LeuLeuTrpArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysVal
5341 ACCTCCTATGGAGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAG
     TGGAGGATACCTCCGTCCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTTC

     ValIleLeuAspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerVal
5401 TGGTGATTCTGGACTCCTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCG
     ACCACTAAGACCTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGC

     ProAlaGluIleLeuArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArg
5461 TACCCGCAGAAATCCTGCGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGC
     ATGGGCGTCTTTAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCG

     ProAspTyrAsnProProLeuValGluThrTrpLysLysProAspTyrGluProProVal
5521 GGCCGGACTATAACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTG
     CCGGCCTGATATTGGGGGGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGAC

     ValHisGlyCysProLeuProProProLysSerProProValProProProArgLysLys
5581 TGGTCCATGGCTGTCCGCTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGA
     ACCAGGTACCGACAGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCT

     ArgThrValValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArg
5641 AGCGGACGGTGGTCCTCACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCA
     TCGCCTGCCACCAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGT

     SerPheGlySerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGlu
5701 GAAGCTTTGGCAGCTCCTCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTG
     CTTCGAAACCGTCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGAC

     ProAlaProSerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProPro
5761 AGCCCGCCCCTTCTGGCTGCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCC
     TCGGGCGGGGAAGACCGACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGG

     LeuGluGlyGluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSer
5821 CCCTGGAGGGGGAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTA
     GGGACCTCCCCCTCGGACCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCAT

     GluAlaAsnAlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeu
```

FIG. 47-6

```
5881 GTGAGGCCAACGCGGAGGATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCAC
     CACTCCGGTTGCGCCTCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTG

      ValThrProCysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeu
5941 TCGTCACCCCGTGCGCCGCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGT
     AGCAGTGGGGCACGCGGCGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCA

      LeuArgHisHisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLys
6001 TGCTACGTCACCACAATTTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGA
     ACGATGCAGTGGTGTTAAACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCT

      LysValThrPheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGlu
6061 AGAAAGTCACATTTGACAGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGG
     TCTTTCAGTGTAAACTGTCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCC

      ValLysAlaAlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSer
6121 AGGTTAAAGCAGCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCA
     TCCAATTTCGTCGCCGCAGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGT

      LeuThrProProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCys
6181 GCCTGACGCCCCCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTT
     CGGACTGCGGGGGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAA

      HisAlaArgLysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsn
6241 GCCATGCCAGAAAGGCCGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACA
     CGGTACGGTCTTTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGT

      ValThrProIleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGlu
6301 ATGTAACACCAATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTG
     TACATTGTGGTTATCTGTGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGAC

      LysGlyGlyArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCys
6361 AGAAGGGGGGTCGTAAGCCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGT
     TCTTCCCCCCAGCATTCGGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACA

      GluLysMetAlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSer
6421 GCGAAAAGATGGCTTTGTACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCT
     CGCTTTTCTACCGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGA

      TyrGlyPheGlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSer
6481 CCTACGGATTCCAATACTCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGT
     GGATGCCTAAGGTTATGAGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCA

      LysLysThrProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGlu
6541 CCAAGAAAACCCCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTG
     GGTTCTTTTGGGGTTACCCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGAC

      SerAspIleArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArg
6601 AGAGCGACATCCGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCC
     TCTCGCTGTAGGCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGG

      ValAlaIleLysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArg
6661 GCGTGGCCATCAAGTCCCTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAA
     CGCACCGGTAGTTCAGGGAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTT

      GlyGluAsnCysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGly
6721 GGGGGAGAACTGCGGCTATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTG
     CCCCCCTCTTGACGCCGATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACAC

      AsnThrLeuThrCysTyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAsp
6781 GTAACACCCTCACTTGCTACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGG
     CATTGTGGGAGTGAACGATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCC

      CysThrMetLeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGln
6841 ACTGCACCATGCTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCC
     TGACGTGGTACGAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGG
```

FIG. 47-7

```
      GluAspAlaAlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProPro
6901  AGGAGGACGCGGCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCC
      TCCTCCTGCGCCGCTCGGACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGG

      GlyAspProProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnVal
6961  CTGGGGACCCCCCACAACCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACG
      GACCCCTGGGGGGTGTTGGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGC

      SerValAlaHisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThr
7021  TGTCAGTCGCCCACGACGGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAA
      ACAGTCAGCGGGTGCTGCCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTT

      ProLeuAlaArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGly
7081  CCCCCCTCGCGAGAGCTGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAG
      GGGGGGAGCGCTCTCGACGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATC

      AsnIleIleMetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePhe
7141  GCAACATAATCATGTTTGCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCT
      CGTTGTATTAGTACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGA

      SerValLeuIleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAla
7201  TTAGCGTCCTTATAGCCAGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGG
      AATCGCAGGAATATCGGTCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCC

      CysTyrSerIleGluProLeuAspLeuProProIleIleGlnArgLeu
7261  CCTGCTACTCCATAGAACCACTTGATCTACCTCCAATCATTCAAAGACTC
      GGACGATGAGGTATCTTGGTGAACTAGATGGAGGTTAGTAAGTTTCTGAG
```

FIG. 47-8